(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 363 176 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
14.02.1996 Bulletin 1996/07

(51) Int Cl.[6]: **G06F 9/312**, G06F 9/315, G06F 9/30

(21) Application number: **89310151.9**

(22) Date of filing: **04.10.1989**

(54) **Word organised data processors**

Prozessoren für wortorganisierte Daten

Processeurs de données organisées en mot

(84) Designated Contracting States:
DE FR GB

(30) Priority: **07.10.1988 US 255096**

(43) Date of publication of application:
11.04.1990 Bulletin 1990/15

(73) Proprietor: **International Business Machines Corporation**
Armonk, N.Y. 10504 (US)

(72) Inventor: **Groves, Randall Dean**
Austin, Texas 78727 (US)

(74) Representative: **Bailey, Geoffrey Alan**
Winchester, Hampshire SO21 2JN (GB)

(56) References cited:
EP-A- 0 130 380
US-A- 3 916 388
US-A- 4 556 951
WO-A-88/04806
US-A- 4 542 456

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 0 363 176 B1

Printed by Jouve, 75001 PARIS (FR)

Description

The present invention relates to a data processing system and more specifically to an apparatus for manipulating character strings in a data processing system.

Data processing systems access and manipulate data. A uniform amount of data is used for transferring and manipulating by the processor. This uniform amount is commonly the amount of data that is stored as a single unit in a memory. In modern data processing systems, this storage unit of data is termed a word and consist of 16 binary digits or bits. Words may also consist of 32, 48 or 64 bits depending on the characteristic of the specific data processing system. It is also desirable to break up words into smaller units such as bytes (8 bits), or nibbles (4 bits).

As data processing systems increase in capability, it is more advantageous to transfer large amounts of data as single units. However, it is also desirable to be able to access small amounts of data, such as bytes or nibbles, in data processing operations. The storage of data units such as words in memory requires the word be the smallest unit of data access. In other words, when one accesses data in memory, one must access a word at a time.

Characters strings are specific character data stored in a contiguous sequence. Commonly, a single character is less than a word in length, usually a byte. Since the length of character strings may vary, storing these variable length character strings only on word boundaries would result in the inefficient use of memory storage, i.e., the number of bytes in a specific character sequence are not always equal to a number of complete words of memory used. Therefore, it is desirable to access character sequences in memory that are not stored on word boundaries.

Since previous data processing systems have used 8 bit, 16 bit or 32 bit word lengths for storage and transfer units, the problem of efficient storage of character strings has not been of great interest. However, more modern data processing systems are increasing the number of bits in a storage or transfer unit. Therefore, the accessibility of bytes within data words is becoming more important.

The following prior art addresses character sequence processing. U. S. Patent No. 4,556,951 entitled, "Central Processor With Instructions For Processing Sequences Of Characters", discusses a data processing system that has been adapted for processing character sequences and more specifically for examining character sequences for a predetermined characteristic.

U. S. Patent No. 4,625,295 entitled, "Textual Comparison System For Locating Desired Character Strings And Delimiter Characters", discloses a data processing system including a text comparison circuit that is used to implement different search strategies for detecting predefined words in a text sequence.

The Advanced Micro Devices publication, Am29000 User's Manual, on page 8-36, lists a compare byte instruction which compares bytes of one register to corresponding bytes of another register.

IBM Technical Disclosure Bulletin, Volume 23, No. 5, October, 1980, pages 2051-2056 entitled "Scan Instruction" discloses an instruction for scanning through a character string from left to right in a byte by byte fashion.

IBM Technical Disclosure Bulletin, Volume 20, No. 8, January, 1978, entitled "Variable Length Byte string Compare and Move Instruction" discloses an instruction for scanning a command string for a delimiter or for the first non-delimiter character.

EP-A-0,130,380 describes a RISC machine having a mechanism for performing full shift, merge, insert and bit alignment functions within one operating machine cycle. Such a mechanism allows data from a source field to be stored in a destination field in main memory beginning at any desired byte address.

The object of the present invention is to provide an apparatus for loading and storing character sequences consisting of data units smaller than the unit of data storage or data processing.

Accordingly the present invention provides a data processing system configured to transfer data in word units between a memory and registers, the system having an apparatus for manipulating contiguous variable length sequences of data stored in data units smaller than a word unit, said apparatus comprising: instruction decoding means for decoding instructions for manipulating a contiguous variable length sequence of the data units, and providing a command signal in response to the manipulation instructions; access means, responsive to a command signal from said instruction decoding means, for accessing said contiguous variable length sequence of data units from said memory even if a first data unit of said sequence is stored in said memory at a location other than a word boundary; and aligning means for arranging said sequence for storage in a plurality of data processing registers, and for providing the arranged sequence to the plurality of processing registers; the system being characterised in that: said aligning means comprises both an alteration means for altering the locations of data units in the sequence, and a comparison means for comparing data units in the sequence with a predefined data unit indicating termination of the sequence of data units; whereby the last unit in the variable length sequence can be identified and the sequence can be aligned such that said first data unit of the sequence is aligned on a first register word boundary.

The present invention provides a word organised data processor having a natural data flow between memory and registers in transfer units of words aligned in both memory and registers on word boundaries, wherein, to enable handling of variable length sequences of information items, each item individually shorter than a word - e.g., character strings - apparatus is inserted between the memory and the registers, responsive to invoked individual instructions, to transfer

data therebetween, word boundary aligned at the registers and variably aligned at memory.

There is disclosed hereinafter, in a data processing system configured to transfer data in word units, an apparatus is provided for manipulating, in response to a single data manipulation instruction, a contiguous variable length sequence of data stored in units smaller than a word unit. The apparatus includes an instruction decoding circuit for decoding instructions and providing a command signal in response to the decoding of the single manipulation instruction. A memory circuit is provided for storing data as addressable words wherein each word includes a plurality of smaller data units. Additional circuitry is provided for accessing a contiguous sequence of the smaller data units in the memory circuit in response to a command signal from the instruction decoding circuit. The first of the smaller data units of the unit sequence stored in the data memory is not stored on a word boundary in the memory circuit. A circuit for providing the smaller data unit sequence to a plurality of data processing registers is also included. This circuit provides that the first smaller data unit in the sequence be aligned with the first registered boundary.

In the preferred embodiment disclosed, circuitry performs the function of fetching character strings from byte locations in memory regardless of whether the bytes are stored on word boundaries. This circuitry includes a byte rotator circuit that arranges the locations of bytes within a single 32 bit word from memory. The byte rotator together with a byte merge circuit are used to merge bytes together into a sequence for storage in the general purpose registers in the data processing system. By providing the ability to rotate and merge bytes, it becomes possible to access bytes that are not stored on word boundaries in memory and provide them on the appropriate boundaries, i.e., register boundaries, to the data processing system. These bytes in sequence, make up character strings which can be manipulated as required in the general purpose registers. A second byte rotator together with a second byte merge capability enables one to store the sequence in memory on non-word boundaries.

Thus the augmented embodiment provides the optional capability to access and manipulate bytes inside of words in a machine that naturally handles words, promoting the efficient storage of character strings in memory by not requiring that the character strings be stored on word boundaries while still enabling individual bytes of the character strings sequence to be manipulated as required by the application running in the data processing system.

The present invention will be described further by way of example with reference to the aforesaid embodiment thereof as illustrated in the accompanying drawings in which:

Figure 1 is a block diagram illustrating the data flow of the embodiment;

Figures 2A through 2D illustrate the data flow for a load string instruction;

Figures 3A through 3D illustrate the data flow for a store string instruction;

Figures 4A through 4D illustrates the data flow for a load string and compare byte instruction;

Figures 5A through 5D illustrate the data flow for a store string instruction for a null terminated character string;

Figure 6 is a flow diagram for the load string instruction;

Figure 7 is a flow diagram for the storage string instruction;

Figure 8 is a flow diagram for the load string and compare byte instruction;

Figure 9A is an illustration of the load string and store string operational code format;

Figure 9B is an example of a character string transfer operation where the character string is specified by byte count;

Figure 9C is an example of instructions for performing a word by word compare operation; and

Figure 9D is an example of a byte comparison operation using the load string and compare byte construction.

The disclosed embodiment adds to a word organised processor an apparatus for fetching and storing character strings. Character strings, as used in the following discussion, include a sequence of characters specified by byte data. The length of the byte sequences can either be specified with a byte count or by a termination character at the end of a sequence.

Figure 1 illustrates in block diagram form the data flow for a character string operation. Referring to Figure 1, the instruction decode unit 10 decodes the string operation instruction and provides the command signal over lines 12 to control logic 14. Control logic 14 includes a fixed point count register 16 (XPCR). Control logic 14 provides the control

signals to operate the different components of the invention such as the byte rotators 22 and 52, registers 26 and 56, byte merge logic 30 and 60 and the byte comparator 34. The control lines from the control logic 14 to these individual units has been omitted for the sake of simplicity. However, the operation of the control logic 14 with these individual components will be explained in flowchart form.

In Figure 1, a byte rotator 22 is shown connected to a register 26 and byte merge logic 30. Byte rotator 22 is connected by bus 18 to a memory 20. In the following discussion, the memory 20 and the term cache will be used interchangeably since this arrangement can be used with both a data cache and a system memory. The memory 20 is addressable by words over the bus 18. The bus is assumed 32 bits wide making the word length 4 bytes. A byte comparator 34 is connected to the register 26 to compare the individual bytes in the register 26 to a predetermine byte. Register 26 is connected to the byte rotator 22 by a bus 24 and also to the byte merge logic 30 by bus 28. Additionally the byte merge logic 30 is connected to receive zeros. The byte rotator 22, register 26, and byte merge logic 30 are one word in length, i.e., 32 bits.

The byte merge logic 30 is connected to the general purpose file 44. These are the general purpose registers for the central processing unit, CPU. The CPU is shown to have access to the general purpose registers file 44 by a bus 48. Each register in the general purpose register file 44 is 32 bits or one word in length. The general purpose file 44 is connected by a bus 50 to a second byte rotator 52. The second byte rotator is also connected to a the register 56 by a bus 54 and byte merge logic 60 by the same bus 54. The register 56 and byte merge logic 60 are connected by a bus 58. The output of the byte merge logic 60 is connected by bus 62 back to memory 20. As before, the capacity of the byte rotator 52, register 56, and byte merge 60 is one word in length.

The control logic 14 functions in the following manner. If the XPCR contains zero, then no operation is performed, otherwise, the following is computed by the control logic 14:

LSBS = 2 Least Significant bits of Starting Address

LSBX = 2 Least Significant bits of (XPCR1 - 1)

Note: All Mask and Mark fields are shown in descending order from Most Significant Byte to Least Significant Byte.

The control logic 14 compute the number of accesses required by the string instructions using the following formula:

$$\text{Number of accesses required} = 1 + (\text{XPCR-1})/4 + (\text{LSBS}+\text{LSBX})/4$$

where '/ ' indicates integer division.

TABLE 1

Load String		
LSBS	# of Bytes to Rotate Left	Load Byte Merge Mask
00	0	0000
01	1	0001
10	2	0011
11	3	0111

TABLE 2

Load String	
LSBX	Last Access Mask Bits
00	1000
01	1100
10	1110
11	1111

The Load Byte Merge Mask controls the Byte Merge 30. A '1 ' indicates that this byte from memory 20 should replace this byte from Register 26. A '0 ' indicates that the byte from Register 26 should be used as is. The control logic 14 generates the Load Byte Merge Mask as shown in Table 1.

The Last Access Mask bits are used to control the Byte Merge 30 for the last access. A '1 ' indicates that this byte should be passed through as controlled by the Load Byte Merge Mask. A '0 ' indicates that a value of binary '00000000' should be passed for this byte. The control logic 14 generates the Last Access Mask bits as shown in Table 2.

TABLE 3

Store String		
LSBS	# of Bytes to Rotate Right	Store Byte Merge Mask
00	0	0000
01	1	1000
10	2	1100
11	3	1110

TABLE 4

Store Strings With Only One Access		
# of bytes in XPCR	LSBS	Memory Mark Bits
1	00	1000
1	01	0100
1	10	0010
1	11	0001
2	00	1100
2	01	0110
2	10	0011
3	00	1110
3	01	0111
4	00	1111

TABLE 5

Store Strings With More Than One Access			
LSBS	First Mark Bits	LSBS+LSBX	Last Mark Bits
00	1111	00	1000
01	0111	01	1100
10	0011	10	1110
11	0001	11	1111

The Store Byte Merge Mask controls the Byte Merge 60. A '1 ' indicates that this byte from register 56 should replace this byte from the General Purpose Register File 44. A '0 ' indicates that the byte from the General Purpose Register File 44 should be used as is. The control logic 14 generates the Store Byte Merge Mask as shown in Table 3.

Mark bits are sent to the Memory 20 with every store access. A '1 ' indicates that this particular byte should be modified in the memory 20. A '0 ' indicates that this byte should remain unchanged in the Memory 20. For Store String instructions which access only one word, only one set of mark bits are generated by the control logic 14 as shown in Table 4. For Store String instructions which access more than one word, the control logic 14 must compute a set of mark bits for both the first and last accesses as shown in Table 5. For all other accesses between the first and last accesses, a set of mark bits containing all ' 1 ' s is used.

To understand the operation hereof, it is best to view an example of a string transfer. In Figure 2A, the contents of memory is illustrated as word 1 and word 2 in memory 20A and word 3 and word 4 in memory 20B. Word 1 and word 2 contain the string"ABCDE". In this example, the string "ABCDE" will be transferred from the word 1 and word 2 location to the word 3 and word 4 location. Presently the word 3 and word 4 contain a second character string "VWXYZ".

When the load string instruction is first received, the address of word 1 is computed and the contents of word 1 is fetched. This is illustrated in Figure 2B by the contents of the memory bus in 18A. Note that the byte count has been stored in the XPCR register 16A. The contents of the word 1 are received and they are rotated left 2 bytes as illustrated by the byte rotate logic 22A. This is stored in register 26A. The byte merge logic 30A and the general purpose register 44A are unchanged.

In Figure 2C the address of word 2 is computed and then word 2 is fetched as illustrated by the memory bus 18B. The contents are then rotated left 2 bytes by the byte rotate logic 22B. A byte merge operation is then performed with the left 2 bytes of register 26B and the right 2 bytes of the byte rotate logic 22B which are output from the byte merge logic 30C as shown. The contents of the byte merge logic in 30C are then stored in the general purpose registers 44C as shown. The byte rotate logic 22B contents are stored into register 26C. Since the byte count indicates that all of the string has been fetched, the final step is to merge the remaining contents of register 26D with the remaining string. Therefore, the left most byte in register 26D is merged in the byte merge logic 30E with trailing zeros and the contents then stored in the general purpose register file 44E as shown. Note that the general purpose register file 44E includes 2 sequential registers for storing the sequence.

The next operation is to store the contents of the general purpose registers to memory in words 3 and words 4 (See Figure 3A). In Figure 3B, the count remains at 5 in the XPCR register 16H. Since the first word in memory is to contain only 3 bytes, the contents of the general purpose register containing the first 4 bytes is rotated right one byte and the memory is told to only update bytes 1, 2 and 3 (see bytes indicated for the memory bus 62H). Bus 62 (Figure 1) includes lines that indicate to the memory 20 which bytes are to be updated.

In Figure 3C, the next word 4 is addressed and the remaining portions of the data string is merged in the byte merge logic 60J and provided onto the memory bus 62J shown modifying the bytes as indicated. the memory 20L illustrates the final transfer of the string (See Figure 3D).

In the preceding example the character string was specified by a count. In the following example, the character string is specified in length by a termination character at the end of the character sequence. However, count register 16 must be loaded with a count larger than the actual string itself. The load string and compare byte instruction is then executed where the byte that is being compared is the terminal character. In Figure 4A, the memory 20 contains the string to be transferred consisting of the character string "ABCDE" and the termination character "0". This character string will be transferred from words 1 and 2 in memory 20M to words 3 and 4 in memory 20N.

In Figure 4B, the fixed point count register 16M is loaded with the number 8 (which is larger than the 6 bytes to be transferred). The address for word 1 is then computed and the contents of word 1 are accessed on memory bus 18M. This contents are rotated two bytes to the left in byte rotator logic 22M and stored in register 26M as shown.

In Figure 4C, word 2 is fetched from memory as shown in memory bus 18N and again is rotated 2 bytes to the left as illustrated in the byte rotate logic 22N. The left 2 bytes of register 26N are merged with the right 2 bytes of the byte rotator logic 22N in the byte merge logic 30P. The byte merge logic contents 30P are stored in the general purpose register 44P. Register 26P now contains the previous contents of the byte rotate logic 22N. The last step merges the contents of register 26Q (left 2 bytes) with the right 2 bytes of the byte rotator logic 22Q in the byte merge logic 30R. The count in the fixed point count register 16Q has been updated since a positive comparison to the termination character was made. The contents of the byte merge logic 30R are then stored into the general purpose registers the general purpose registers as shown in 44R.

Again, Figures 5A through 5D illustrate the storage of the general purpose register containing the string, terminated by the null character, to the memory location 20N. In Figure 5B, the contents of the general purpose register 44S containing the first 4 bytes is rotated 1 byte to the right. This contents is then placed in register 56S to the byte merge logic 60S. The 3 bytes indicated on the memory bus 62S are updated.

In Figure 5C, the contents of register 56T are then merged with the remaining sequence in the byte rotate logic 52T to form the contents of the byte merge logic 60U which is placed on the memory bus 62T to update the bytes as indicated. This results in the transfer of the character string to the memory location 20W in Figure 5D.

Figures 6 illustrates a flowchart for the load string instruction. The load string instruction is first decoded in Step 100. In Step 102, the starting byte address is computed. This includes accessing the cache or memory word that contains that byte address. In Step 104, determination is made as to whether the string has been terminated. If not, in Step 106 the starting byte is left justified and stored in the register 26, the address is incremented and the next word is fetched. In Step 110, again it is determined whether the end of the string has been encountered. If not, in Step 112, the byte rotate operation is performed and the newly fetched bytes are merged with the previously fetched bytes in the register 26. The merge contents is then stored to the general purpose register 44. The rotated contents are then saved in the register 26. The general purpose register pointer 46 is incremented, the address is then incremented and the next word fetched.

Returning to Step 108, the end of the string has been accessed, the starting byte is left justified and the last bytes merged with trailing zeros. The contents are then stored in the general purpose register. In Step 117, the next instruction is decoded. Returning to Step 114, again when the end of string has been accessed, the contents of the memory bus are rotated and merged with trailing zeroes and the contents in the merge logic are stored in the general purpose register. The process then proceeds to the next instruction in Step 116.

Figure 7 illustrates the store string instruction. The store string instruction is decoded in Step 120. In Step 122, the starting byte address is then computed. From this it is determined the bytes that will be stored in the modified memory word. These data is fetched from the general purpose registers 44 and rotated to align the bytes in their appropriate

positions. These bytes are sent to memory with an indication to memory to modify the specific bytes. The rotated data is saved in the register 56. In Step 124, it is determined whether the end of the string is being encountered. If not, in Step 126, the address for the next word is incremented, the general purpose pointer is incremented, and the number of bytes to be updated is then computed. Data is fetched from the general purpose registers 44, rotated and merged and provided to memory. The rotated contents is then saved in the register 56. When the end of the string is encountered, the process proceeds to Step 128 to decode the next instruction.

Figure 8 illustrates the load string and compare byte instruction. In Step 130, load string and compare byte instruction is decoded. In Step 132, the starting address of the byte is computed and the word containing this byte is accessed. In Step 134, it is determined whether the end of the string has been encountered. If not, in Step 136, it is determined if any of the bytes are equivalent to the comparison byte. If not, in Step 138, the starting byte is left justified and then saved. The address is incremented and the next word is accessed. The byte rotate and merge function are performed and the contents are stored in the general purpose registers 44. The rotated contents are then saved in register 26. In Step 146, it is again determined whether the end of the string has been accessed. If not, in Step 148, it is determined if any of the bytes are equivalent to the comparison byte. If not, in Step 150, the general purpose register pointer 46 is incremented, the byte address is incremented and the next word is accessed from memory. The byte rotate and merge operations are performed with contents stored in the general purpose registers 44. The rotated contents are stored register 26. Returning to Step 140, the starting byte is left justified and stored in the the general purpose registers 44. In Step 142, it is determined if any of the bytes are equivalent to the comparison byte. If so, the process computes the number of the first matching byte and then updates the byte count and flag accordingly. The process proceeds to the Step 159, to decode the next instructions.

Returning to Step 152, the general purpose register pointer 46 is incremented and the byte rotate and merge operations are performed with contents stored in the general purpose registers 44. In Step 154, it is determined whether any of the bytes are equivalent to the comparison byte. If so, in Step 156, the number of the first matching byte is computed, the byte count changed and flag set. In Step 158, the next instruction is decoded.

Figure 9A illustrates the format for the load string and store string instruction. These string instructions can be used to perform rapid string moves by being coded in a small loop. The address of the starting byte of the string is computed by adding general purpose register RA + RB. For load string "ls", the string is loaded starting with register RT, then RT+1, and so on until the end of the string. For the store string instruction "ss", the string is stored starting with register RT, then RT+1, and so on until the end of the string. The length of the string is contained in the register called the Fixed Point Control Register (XPCR). For example, if the XPCR is loaded with the count of 32 bytes, a counter can be established that represents the number of 32 byte blocks which will be moved and be placed in register Rcnt. The starting address of the source string is in Rs and the starting address of the destination string is in register Rd. Eight working registers are required for 32 bytes (Rw-Rw+7). Register Rx contains the current index value starting at zero. The loop for moving the string is illustrated in Figure 9B.

Because the ls (load string) instruction pads the registers with zeroes past the end of the string, this instruction can be used to perform string comparisons in which the comparison occurs a word at a time. If Rs1 and Rs2 are the starting addresses of the 2 strings and Rv is another set of working registers, the loop illustrated in Figure 9C will perform this function.

For null terminated strings (as in the language C), the Load String and Compare Bytes "lscb" instructions can be used. The byte to be compared with is in a different field within the XPCR. Since lscb instructions update the byte content field in the XPCR to the number of the matching bytes, the store string instruction which follows it will only store up to and including the null character. The register usage is as described above except that Rcnt is not needed since the length of the string is determined by the location of the null character. Assuming the XCPR is loaded with the null character and the byte count is 32, the loop to accomplish the function is illustrated in Figure 9D.

Claims

1. A data processing system configured to transfer data in word units between a memory (20) and registers (44), the system having an apparatus for manipulating contiguous variable length sequences of data stored in data units smaller than a word unit, said apparatus comprising:

   instruction decoding means (10) for decoding instructions for manipulating a contiguous variable length sequence of the data units, and providing a command signal in response to the manipulation instructions;

   access means (14), responsive to a command signal from said instruction decoding means (10), for accessing said contiguous variable length sequence of data units from said memory (20) even if a first data unit of said sequence is stored in said memory (20) at a location other than a word boundary; and

aligning means for arranging said sequence for storage in a plurality of data processing registers (44), and for providing the arranged sequence to the plurality of processing registers (44);

the system being characterised in that:

said aligning means comprises both an alteration means (22, 26, 30) for altering the locations of data units in the sequence, and a comparison means (34) for comparing data units in the sequence with a predefined data unit indicating termination of the sequence of data units;

whereby the last unit in the variable length sequence can be identified and the sequence can be aligned such that said first data unit of the sequence is aligned on a first register word boundary.

2. A system as claimed in Claim 1, further characterised by:
transfer means (52, 56, 60) for transferring the sequence as stored in the data processing registers (44) to the memory (20) and storing that sequence in the memory (20) by updating individual data unit locations in said memory (20).

3. A system as claimed in Claim 2, wherein said transfer means (52, 56, 60) includes means for determining a predetermined data unit in the registers to be a trailing data unit in the last word of the sequence stored in the register.

4. A system as claimed in any of claims 1 to 3, wherein the alteration means (22, 26, 30) comprises a rotator circuit and a merge circuit to alter the location of data units in the sequence for storage in the plurality of data processing registers (44).

5. A system as claimed in any preceding claim, wherein each data unit is the same integral sub-multiple of the processor word size.

6. A system as claimed in any preceding claim, wherein the word units contain a plurality of bytes and the data units are individual bytes.

7. A system as claimed in any preceding claim, wherein the manipulation instruction specifies a first register of a sequence of the registers (44) and the apparatus includes means for storing the sequence in the specified first register and sufficient of the remaining registers in the natural processor nominated sequence thereof.

Patentansprüche

1. Ein Datenverarbeitungssystem, das konfiguriert ist, um Daten in Worteinheiten zwischen einem Speicher (20) und Registern (44) zu übertragen, wobei das System ein Gerät zur Verarbeitung von zusammenhängenden Datenfolgen mit variabler Lenge hat, die in Dateneinheiten gespeichert sind, die kleiner als eine Worteinheit sind, wobei das Gerät enthält:

Befehlsdecodiermittel (10) zum Decodieren von Befehlen zwecks Verarbeitung einer zusammenhängenden Folge von Dateneinheiten mit variablen Längen und die Bereitstellung eines Befehlssignals als Reaktion auf die Verarbeitungsbefehle;

Zugriffsmittel (14), die auf ein Befehlssignal von diesen Befehlsdecodiermitteln (10) reagieren, um von diesem Speicher (20) aus auf die zusammenhängende Folge von Dateneinheiten variabler Länge zuzugreifen, selbst wenn eine erste Dateneinheit aus dieser Folge im Speicher (20) an einer anderen Stelle als an einer Wortgrenze gespeichert ist; und

Ausrichtungsmittel, um diese Folge zur Speicherung in einer Vielzahl von Datenverarbeitungsregistern (44) anzuordnen und zur Bereitstellung der so angeordenten Folge an die Vielzahl von Verarbeitungsregistern (44);

wobei das System dadurch gekennzeichnet wird, daß

die Ausrichtungsmittel sowohl ein Änderungsmittel (22, 26, 30) enthalten, um die Stellen von Dateneinheiten in der Folge zu ändern, und ein Vergleichsmittel (34), um Dateneinheiten in der Folge mit einer zuvor definierten

Dateneinheit zu vergleichen, welche den Abschluß von Dateneinheiten angibt;

wobei die letzte Einheit in der Folge mit variabler Länge identifiziert werden kann und die Folge so ausgerichtet werden kann, daß die erste Dateneinheit in der Folge mit einer ersten Registerwortgrenze ausgerichtet wird.

2. Ein System wie in Anspruch 1 angemeldet, das außerdem gekennzeichnet wird durch Übertragungsmittel (52, 56, 60) zur Übertragung der Zeichenfolge, wie diese in den Datenverarbeitungsregistern (44) in Speicher (20) gespeichert sind, und zur Speicherung dieser Zeichenfolge im Speicher (20), indem die Speicherstellen der einzelnen Dateneinheiten in diesem Speicher (20) aktualisiert werden.

3. Ein System wie in Anspruch 2 angemeldet, wobei diese Übertragungsmittel (52, 56, 60) Mittel enthalten, um eine zuvor bestimmte Dateneinheit in den Registern als nachfolgende Dateneinheit im letzten Wort der in dem Register gespeicherten Zeichenfolge zu bestimmen.

4. Ein System wie in irgendeinem der Ansprüche 1 bis 3 angemeldet, wobei die Änderungsmittel (22, 26, 30) einen Rotator- und einen Mischschaltkreis enthalten, um die Speicherstelle der Dateneinheiten in der Zeichenfolge zur Speicherung in der Vielzahl von Datenverarbeitungsregistern (44) zu ändern.

5. Ein System wie in irgendeinem vorhergehenden Anspruch angemeldet, wobei jede Dateneinheit die gleiche ganze Sub-Multiple des Wortformats des Prozessor ist.

6. Ein System wie in irgendeinem vorhergehenden Anspruch angemeldet, wobei die Worteinheiten eine Vielzahl von Bytes enthalten, und die Dateneinheiten individuelle Bytes sind.

7. Ein System wie in irgendeinem vorhergehenden Anspruch angemeldet, wobei der Verarbeitungsbefehl ein erstes Register aus einer Folge von Registern (44) angibt, und das Gerät Mittel enthält, um die Zeichenfolge in dem angegebenen ersten Register und genügend von den restlichen Registern in der von dem natürlichen Prozessor ausgewählten Zeichenfolge zu speichern.

Revendications

1. Système de traitement de données, configuré pour transférer des données en unités de mot, entre une mémoire (20) et des registres (44), le système ayant un dispositif destiné à la manipulation de séquences de données contigues, de longueur variable, stockées dans des unités de données de taille inférieure à une unité de mot, ledit appareil comprenant :

des moyens de décodage d'instruction (10) destinés à décoder des instructions destinées à manipuler une séquence d'une unité de données contigue de longueur variable et à produire un signal de commande en réponse aux instructions de manipulation;

des moyens d'accès (14), réagissant à un signal de commande venant desdits moyens de décodage d'instruction (10), pour permettre l'accès à ladite séquence d'unité de données contiguë de longueur variable venant de ladite mémoire (20), même si une première unité de données de ladite séquence est stockée dans ladite mémoire (20), en un emplacement autre qu'une limite de mot; et

des moyens d'alignement, pour agencer ladite séquence pour permettre le stockage en une pluralité de registres de traitement de données (44) et pour fournir ladite séquence agencée à la pluralité de registres de traitement (44);

le système étant caractérisé en ce que :

lesdits moyens d'alignement comprennent des moyens de modification (22, 26, 30), destinés à modifier des emplacements d'unités de données dans la séquence, et un moyen de comparaison (34), destiné à comparer les unités de données dans la séquence avec une unité de données prédéfinie, indiquant l'achèvement de la séquence d'unité de données;

de manière que la dernière unité dans la séquence de longueur variable puisse être identifiée et que la séquence

puisse être alignée de façon que ladite première unité de données de la séquence soit alignée sur une première limite de mot de registre.

2. Système selon la revendication 1, caractérisé en outre par :
des moyens de transfert (52, 56, 60) pour assurer le transfert de la séquence, une fois stockée dans les registres de traitement de données (44), à la mémoire (20) et stocker cette séquence dans la mémoire (20) par mise à jour d'emplacements d'unités de données individuels dans ladite mémoire (20).

3. Système selon la revendication 2, dans lequel lesdits moyens de transfert (52, 56, 60) comprennent des moyens pour déterminer une unité de données prédéterminée dans les registres, pour la placer en unité de données de queue dans le dernier mot de la séquence stockée dans le registre.

4. Système selon l'une quelconque des revendications 1 à 3, dans lequel les moyens de modifications (22, 26, 30) comprennent un circuit permutation circulaire et un circuit de fusion, pour modifier l'emplacement des unités de données dans la séquence, pour le stockage dans la pluralité de registres de traitement de données (44).

5. Système selon l'une quelconque des revendications précédentes, dans lequel chaque unité de données est le même sous-multiple entier de la taille de mot du processeur.

6. Système selon l'une quelconque des revendications précédentes, dans lequel les unités de mot comprennent une pluralité de multiplets et les unités de données sont des multiplets individuels.

7. Système selon l'une quelconque des revendications précédentes dans lequel l'instruction de manipulation spécifie un premier registre dans la séquence des registres (44) et le dispositif comprend des moyens pour stocker la séquence dans le premier registre spécifié et dans un nombre suffisant, parmi les registres restants, dans la séquence naturelle énoncée du processeur.

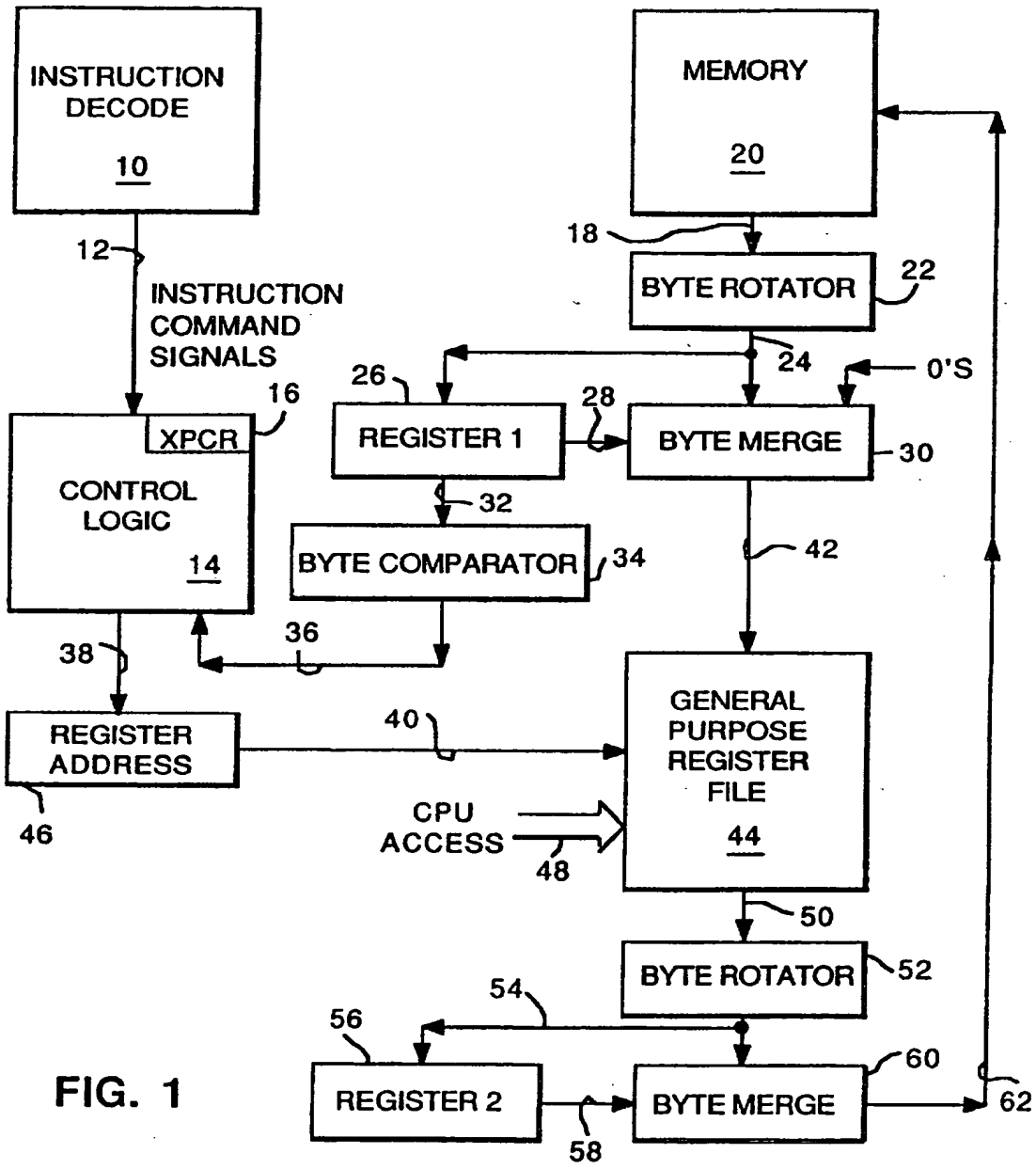

INSTRUCTION DECODE
10
12
INSTRUCTION COMMAND SIGNALS
16
XPCR
CONTROL LOGIC
14
MEMORY
20
18
BYTE ROTATOR
22
24
0'S
26
REGISTER 1
28
BYTE MERGE
30
32
BYTE COMPARATOR
34
42
36
38
REGISTER ADDRESS
40
46
GENERAL PURPOSE REGISTER FILE
44
CPU ACCESS
48
50
BYTE ROTATOR
52
54
56
REGISTER 2
58
BYTE MERGE
60
62

FIG. 1

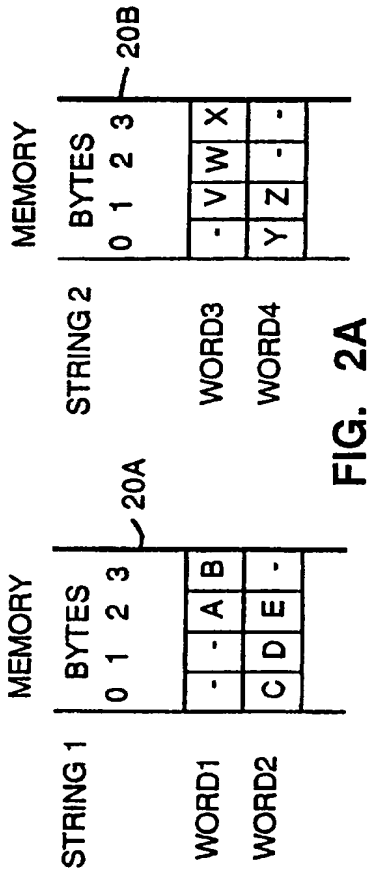

MEMORY
BYTES
0 1 2 3
20A
STRING 1
WORD1
WORD2
- - A B
C D E -
MEMORY
BYTES
0 1 2 3
20B
STRING 2
WORD3
WORD4
- V W X
Y Z - -

FIG. 2A

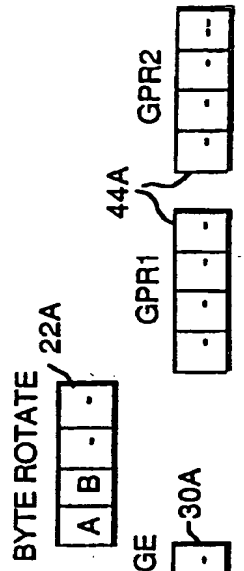

BYTE ROTATE
22A
A B - -
GPR1
44A
GPR2
- - - -
- - - -
BYTE MERGE
30A
- - - -
REGISTER 1
A B - -
26A
XPCR
16A
- - 0 5
(BYTE COUNT)
MEMORY BUS
18A
- - A B

FIG. 2B

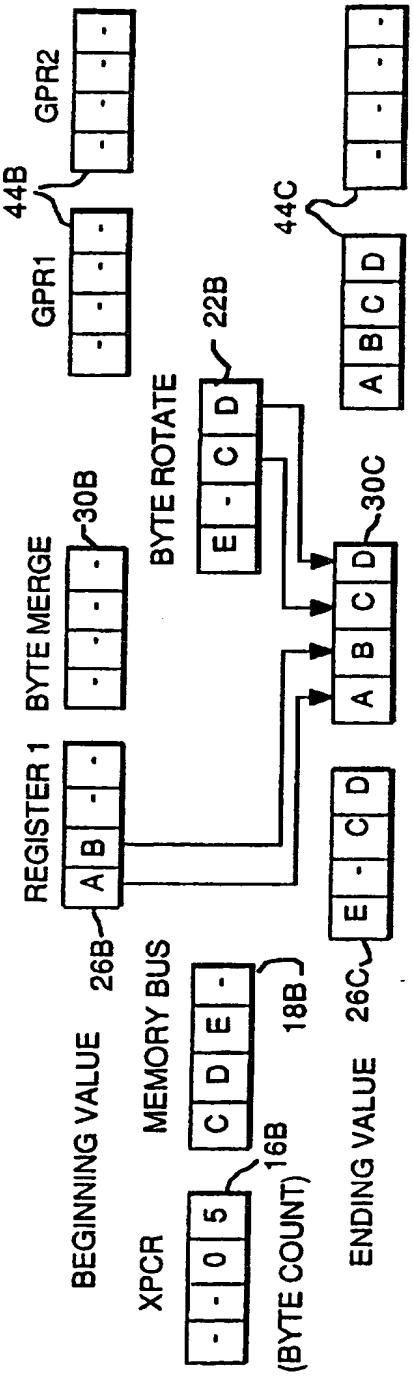

GPR1
44B
GPR2
BYTE MERGE
30B
BYTE ROTATE
22B
44C
REGISTER 1
A B - -
26B
BEGINNING VALUE
XPCR
- - 0 5
(BYTE COUNT) 16B
MEMORY BUS
C D E -
18B
ENDING VALUE
26C
E - C D
A B C D
30C
A B C D

FIG. 2C

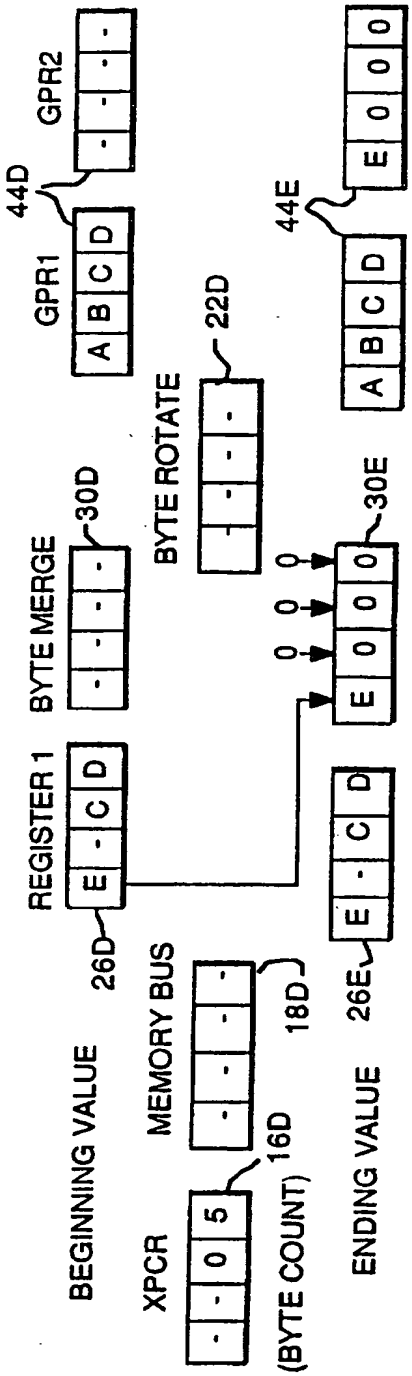

GPR1
44D
GPR2
BYTE MERGE
30D
BYTE ROTATE
22D
44E
REGISTER 1
E - C D
26D
BEGINNING VALUE
XPCR
- - 0 5
(BYTE COUNT) 16D
MEMORY BUS
18D
ENDING VALUE
26E
E - C D
0 0 0 0
30E
A B C D
E 0 0 0

FIG. 2D

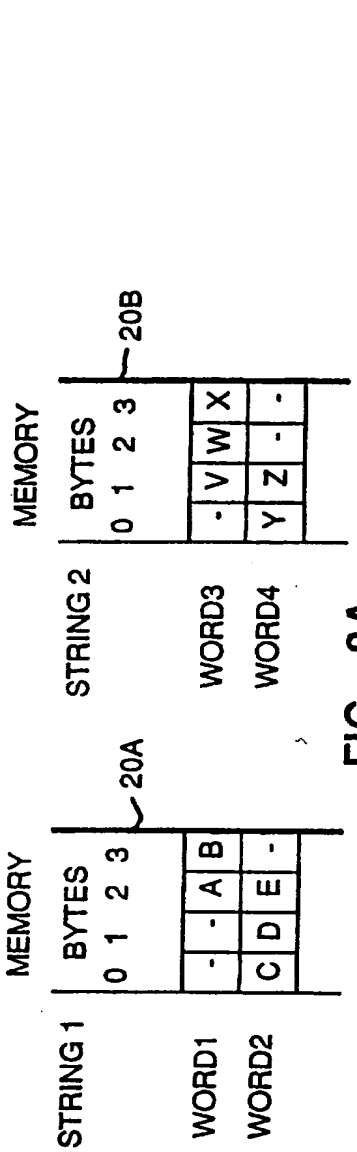

MEMORY
BYTES
0 1 2 3
20A
STRING 1
WORD1
WORD2
- - A B
C D E -
STRING 2
WORD3
WORD4
MEMORY
BYTES
0 1 2 3
20B
- V W X
Y Z - -

FIG. 3A

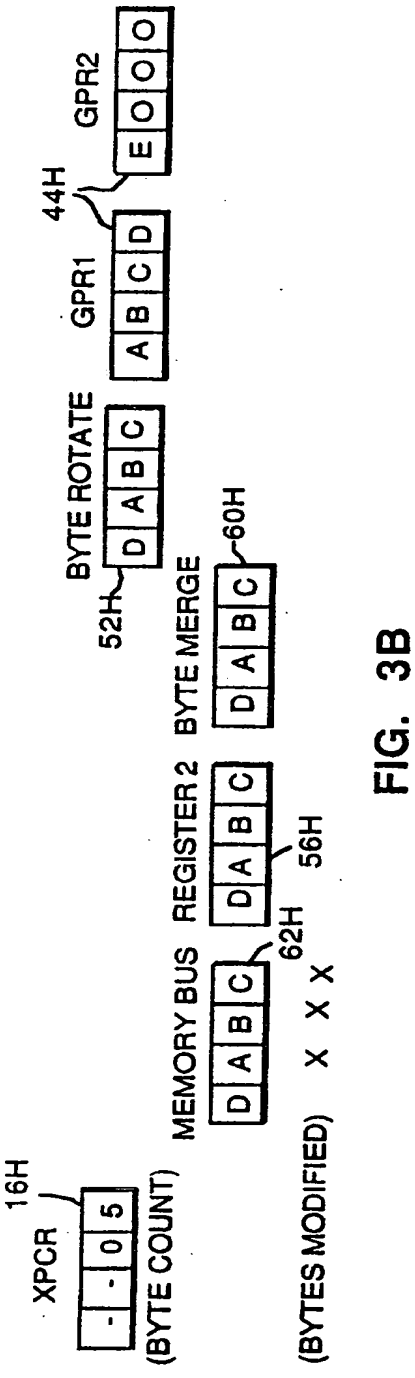

GPR2
E O O O
44H
GPR1
A B C D
BYTE ROTATE
52H
D A B C
BYTE MERGE
60H
D A B C
REGISTER 2
56H
D A B C
MEMORY BUS
62H
D A B C
(BYTES MODIFIED) X X X
XPCR
16H
- - 0 5
(BYTE COUNT)

FIG. 3B

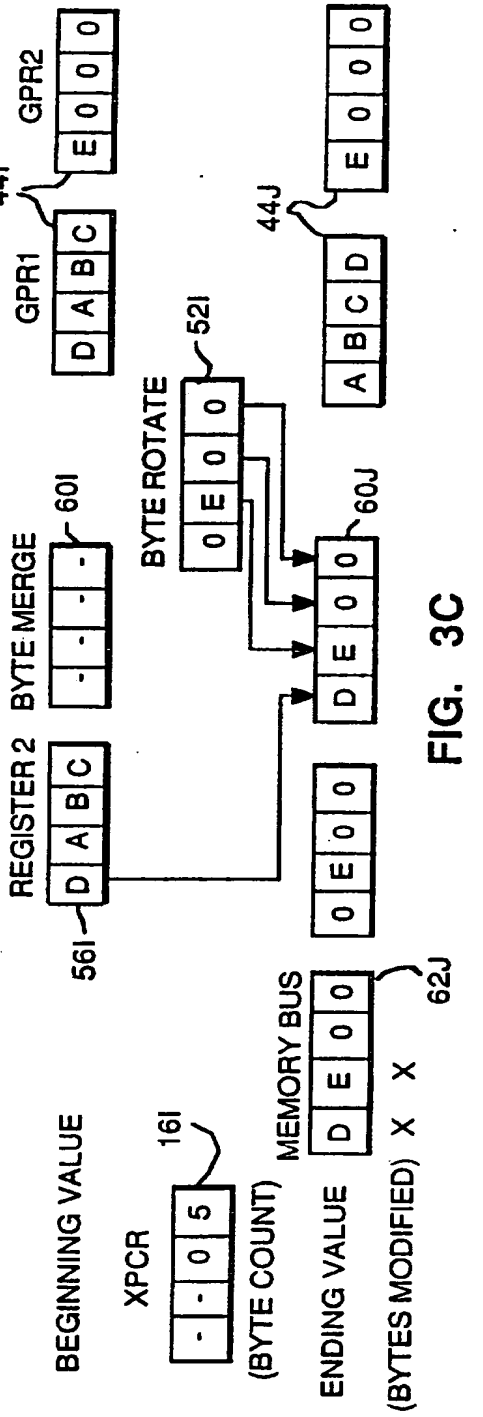
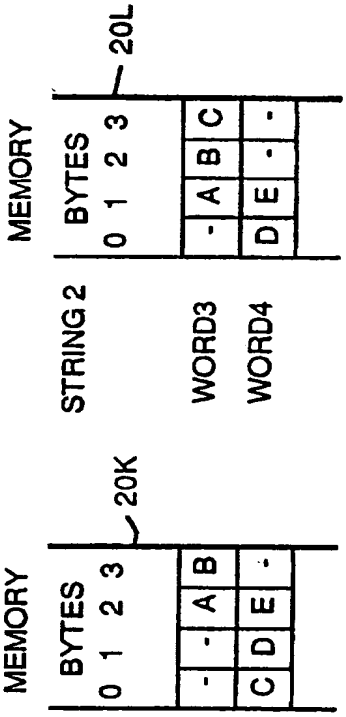

GPR1
D A B C
441
GPR2
E 0 0 0
BYTE MERGE
- - - -
601
BYTE ROTATE
0 E 0 0
521
44J
A B C D
E 0 0 0
REGISTER 2
D A B C
561
60J
D E 0 0
BEGINNING VALUE
XPCR
- - 0 5
161
(BYTE COUNT)
MEMORY BUS
D E 0 0
ENDING VALUE
(BYTES MODIFIED) X X
0 E 0 0
62J

FIG. 3C


MEMORY
BYTES
0 1 2 3
20L
STRING 2
WORD3
- A B C
WORD4
D E - -

MEMORY
BYTES
0 1 2 3
20K
STRING 1
WORD1
- - A B
WORD2
C D E -

FIG. 3D

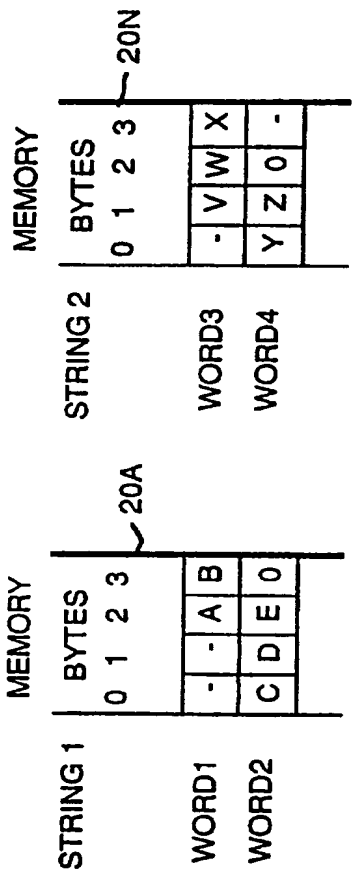

MEMORY
BYTES
0 1 2 3
20A
STRING 1
WORD1
WORD2
- - A B
C D E 0
STRING 2
WORD3
WORD4
- V W X
Y Z 0 -
20N

FIG. 4A

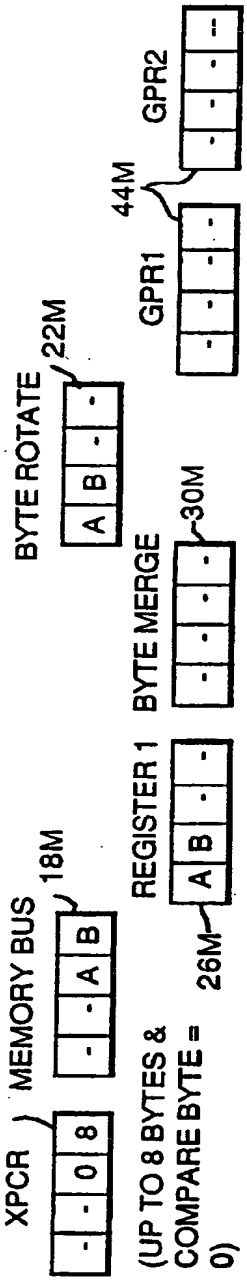

BYTE ROTATE
22M
A B - -
GPR1
44M
GPR2
- - - -
- - - -
- - - --
BYTE MERGE
30M
- - - -
REGISTER 1
A B - -
26M
(UP TO 8 BYTES & COMPARE BYTE = 0)
XPCR
16M
- - 0 8
MEMORY BUS
18M
- - A B

FIG. 4B

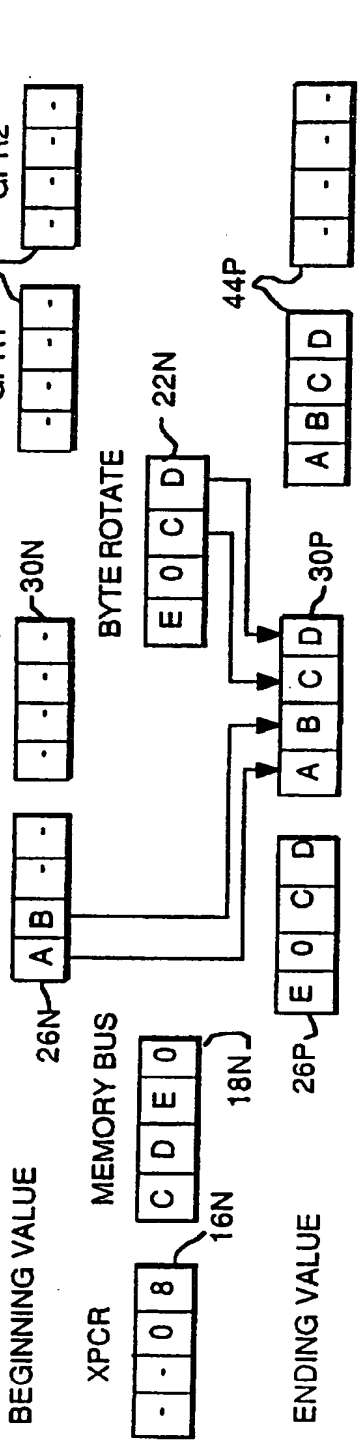

BEGINNING VALUE
XPCR
MEMORY BUS
REGISTER 1
BYTE MERGE
BYTE ROTATE
GPR1
GPR2
ENDING VALUE
26N
16N
18N
30N
22N
44N
26P
30P
44P

FIG. 4C

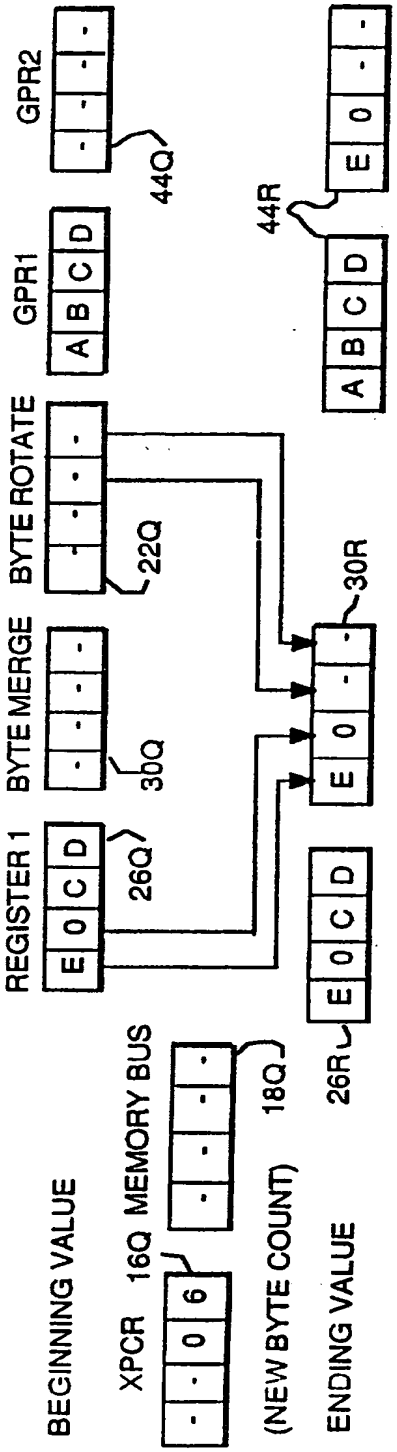

BEGINNING VALUE
XPCR
16Q MEMORY BUS
(NEW BYTE COUNT)
18Q
REGISTER 1
BYTE MERGE
BYTE ROTATE
GPR1
GPR2
ENDING VALUE
26Q
30Q
22Q
44Q
26R
30R
44R

FIG. 4D

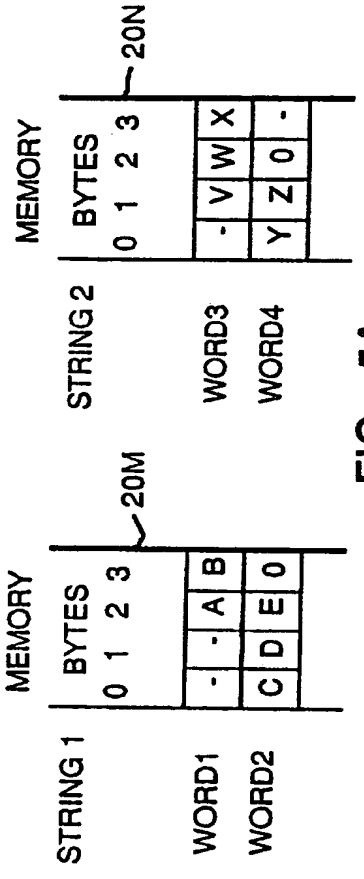

MEMORY
BYTES
0 1 2 3
20M
STRING 1
WORD1
WORD2
- - A B
C D E 0
STRING 2
WORD3
WORD4
20N
- V W X
Y Z 0 -

FIG. 5A

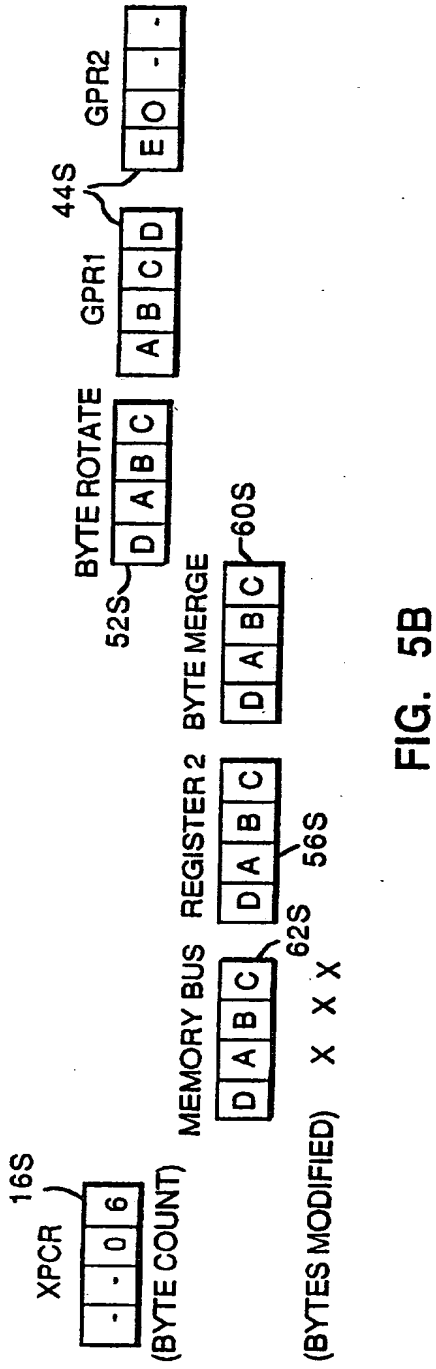

GPR2
E O - -
44S
GPR1
A B C D
BYTE ROTATE
52S
D A B C
BYTE MERGE
60S
D A B C
REGISTER 2
56S
D A B C
62S
MEMORY BUS
D A B C
(BYTES MODIFIED) X X X
16S
XPCR
- - 0 6
(BYTE COUNT)

FIG. 5B

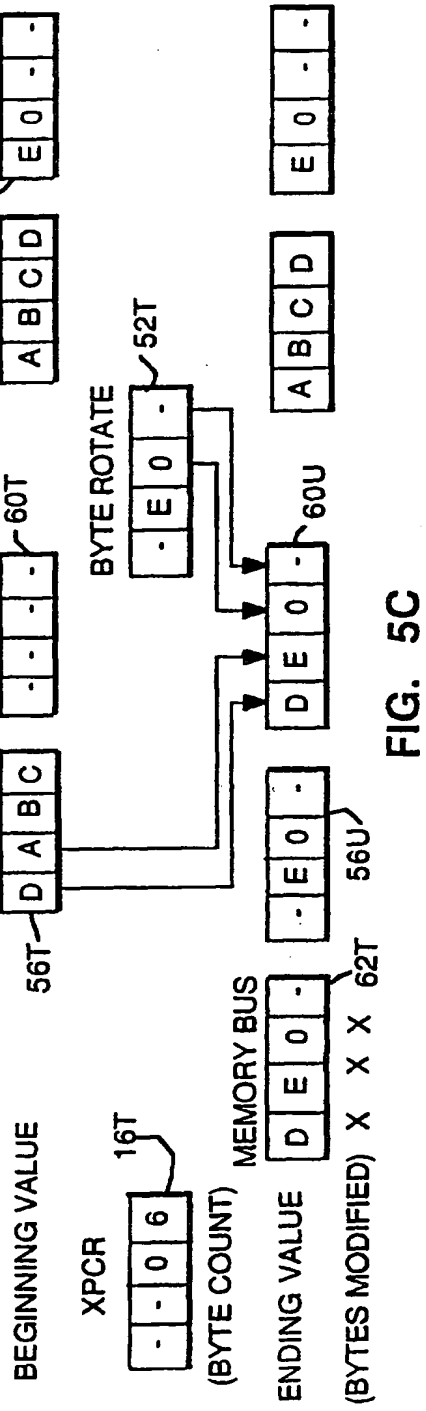

REGISTER 2
BYTE MERGE
BYTE ROTATE
GPR1
GPR2
44T
56T
60T
52T
60U
56U
62T
BEGINNING VALUE
XPCR
16T
(BYTE COUNT)
ENDING VALUE
MEMORY BUS
(BYTES MODIFIED) X X X

FIG. 5C

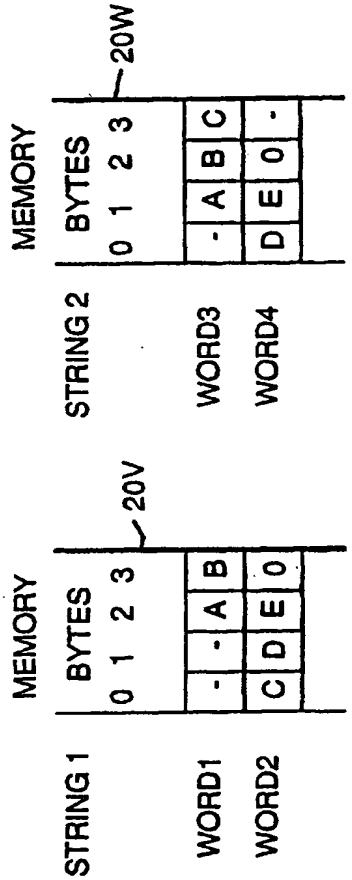

MEMORY
BYTES
0 1 2 3
20W
STRING 2
WORD3
WORD4
MEMORY
BYTES
0 1 2 3
20V
STRING 1
WORD1
WORD2

FIG. 5D

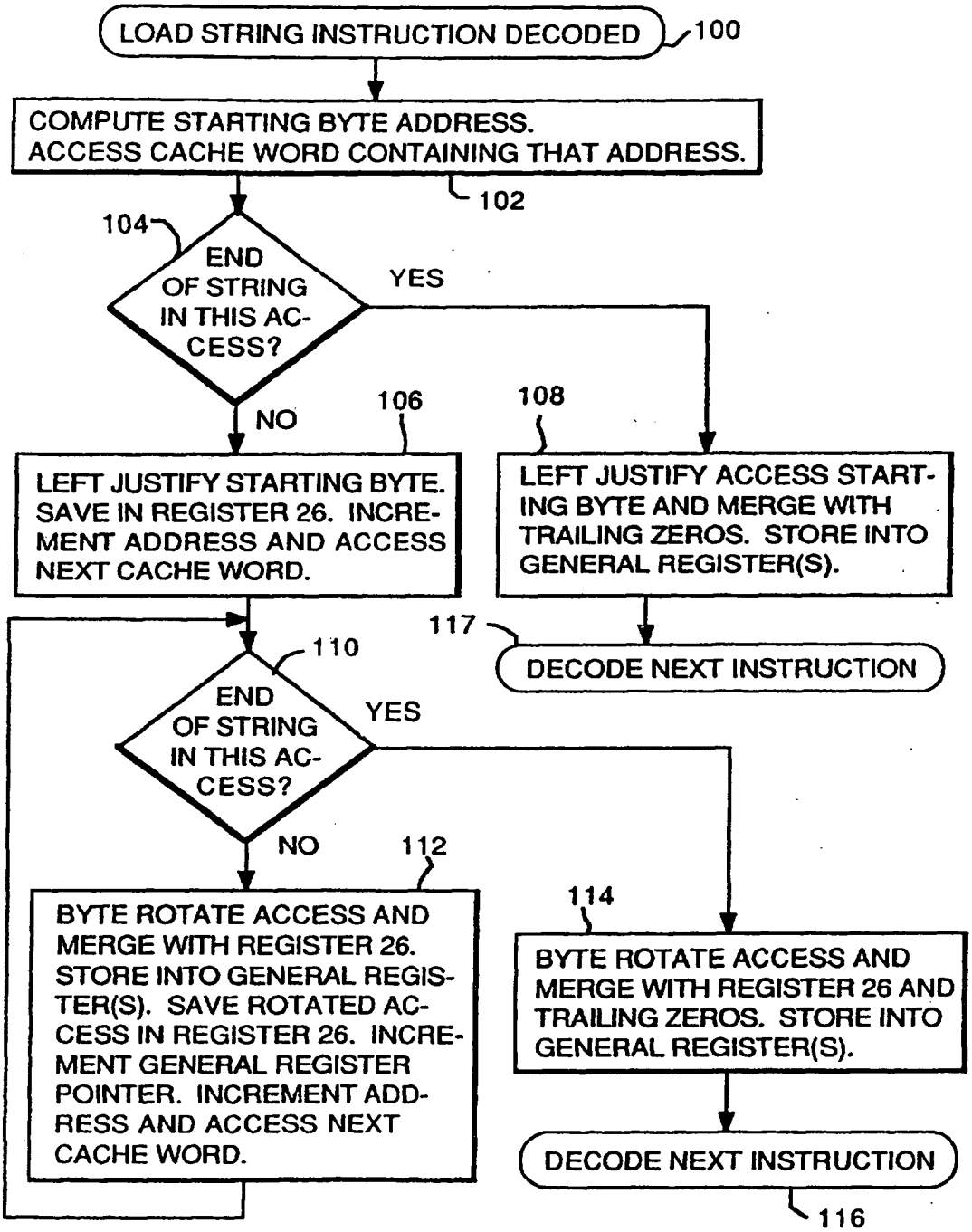

LOAD STRING INSTRUCTION DECODED
100
COMPUTE STARTING BYTE ADDRESS. ACCESS CACHE WORD CONTAINING THAT ADDRESS.
102
104
END OF STRING IN THIS ACCESS?
YES
NO
106
LEFT JUSTIFY STARTING BYTE. SAVE IN REGISTER 26. INCREMENT ADDRESS AND ACCESS NEXT CACHE WORD.
108
LEFT JUSTIFY ACCESS STARTING BYTE AND MERGE WITH TRAILING ZEROS. STORE INTO GENERAL REGISTER(S).
117
DECODE NEXT INSTRUCTION
110
END OF STRING IN THIS ACCESS?
YES
NO
112
BYTE ROTATE ACCESS AND MERGE WITH REGISTER 26. STORE INTO GENERAL REGISTER(S). SAVE ROTATED ACCESS IN REGISTER 26. INCREMENT GENERAL REGISTER POINTER. INCREMENT ADDRESS AND ACCESS NEXT CACHE WORD.
114
BYTE ROTATE ACCESS AND MERGE WITH REGISTER 26 AND TRAILING ZEROS. STORE INTO GENERAL REGISTER(S).
DECODE NEXT INSTRUCTION
116

FIG. 6

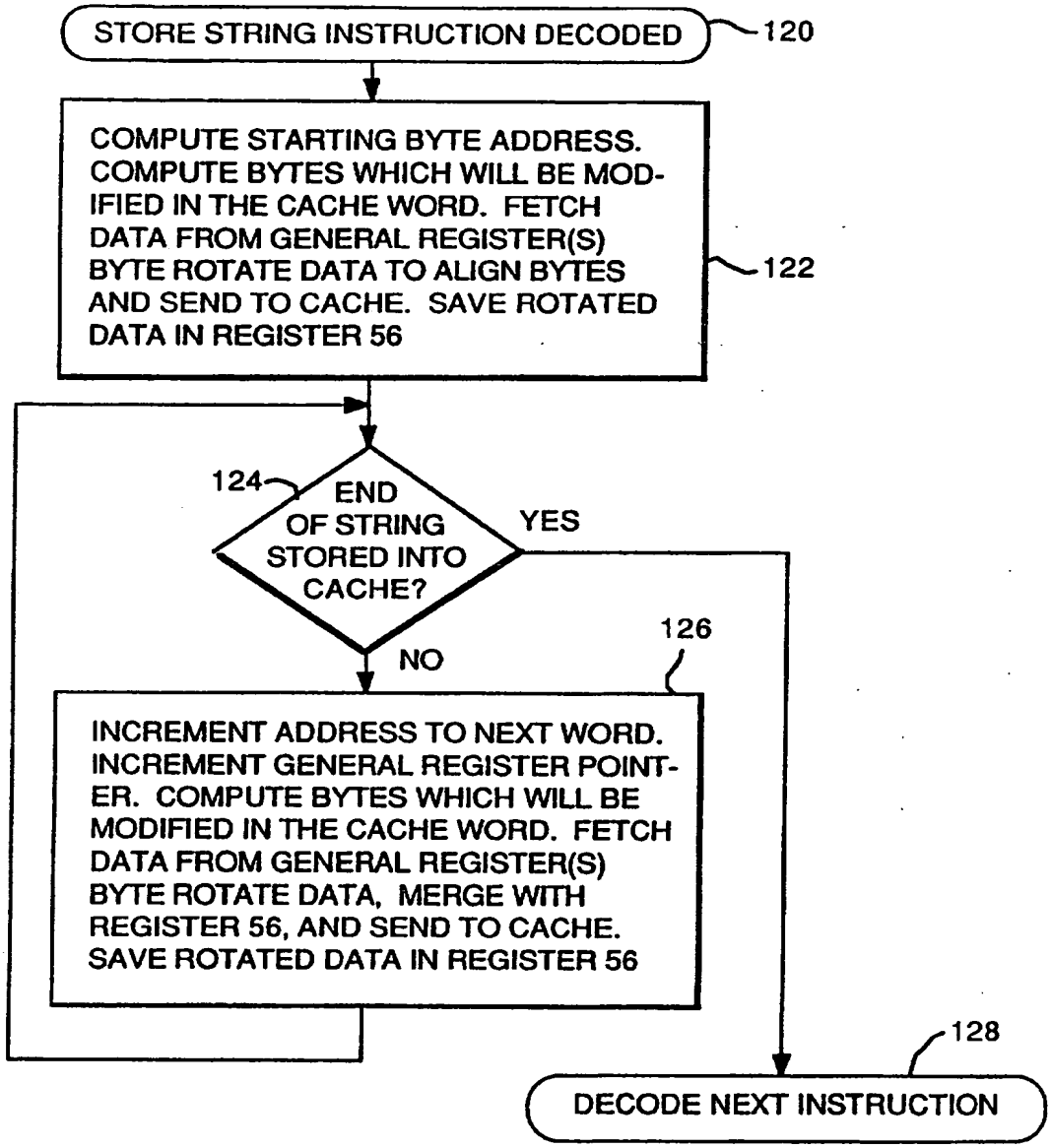

STORE STRING INSTRUCTION DECODED
120
COMPUTE STARTING BYTE ADDRESS. COMPUTE BYTES WHICH WILL BE MODIFIED IN THE CACHE WORD. FETCH DATA FROM GENERAL REGISTER(S) BYTE ROTATE DATA TO ALIGN BYTES AND SEND TO CACHE. SAVE ROTATED DATA IN REGISTER 56
122
124
END OF STRING STORED INTO CACHE?
YES
NO
126
INCREMENT ADDRESS TO NEXT WORD. INCREMENT GENERAL REGISTER POINTER. COMPUTE BYTES WHICH WILL BE MODIFIED IN THE CACHE WORD. FETCH DATA FROM GENERAL REGISTER(S) BYTE ROTATE DATA, MERGE WITH REGISTER 56, AND SEND TO CACHE. SAVE ROTATED DATA IN REGISTER 56
128
DECODE NEXT INSTRUCTION

FIG. 7

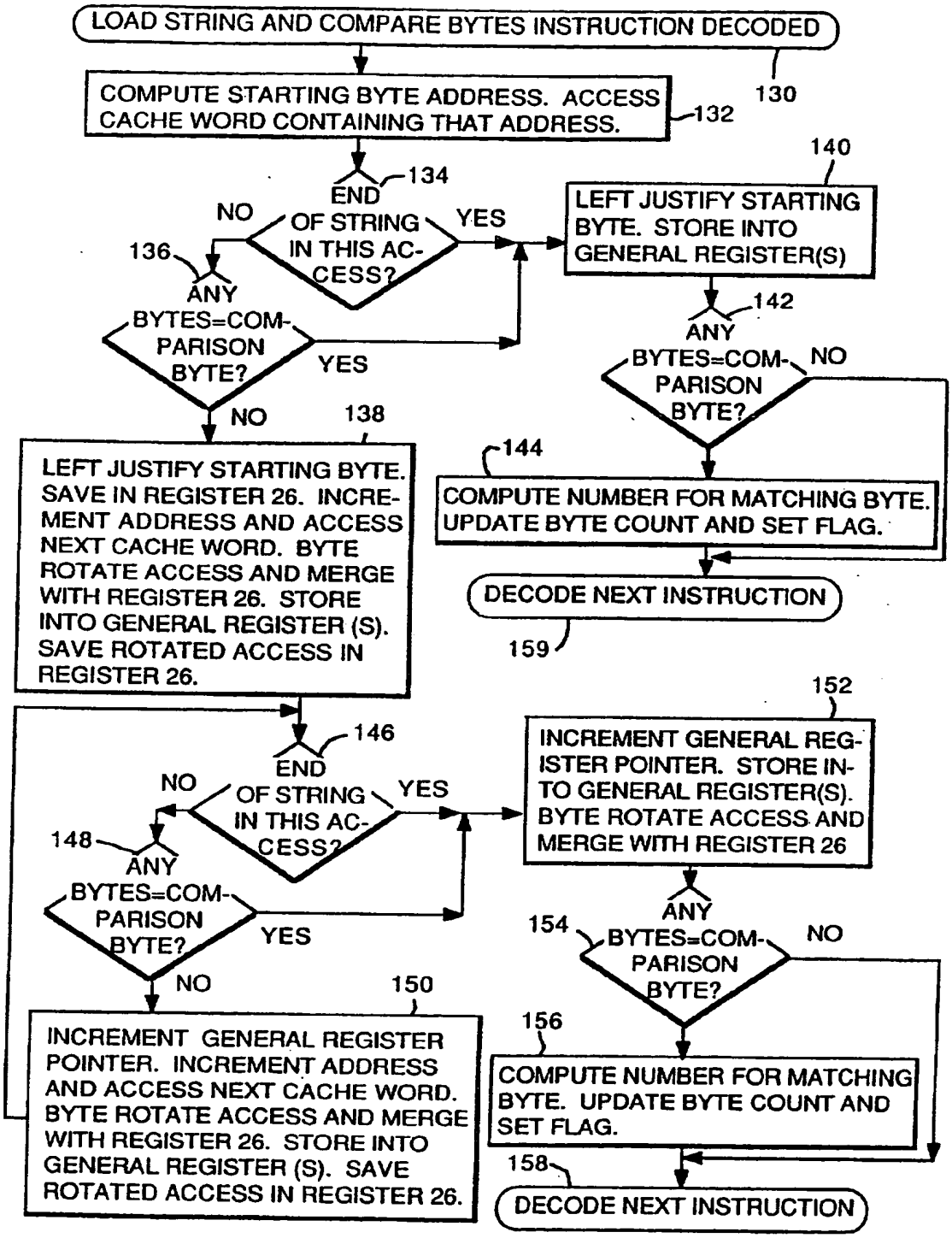

LOAD STRING AND COMPARE BYTES INSTRUCTION DECODED
130
COMPUTE STARTING BYTE ADDRESS. ACCESS CACHE WORD CONTAINING THAT ADDRESS.
132
END OF STRING IN THIS ACCESS?
134
NO
YES
136
ANY BYTES=COMPARISON BYTE?
YES
NO
138
LEFT JUSTIFY STARTING BYTE. SAVE IN REGISTER 26. INCREMENT ADDRESS AND ACCESS NEXT CACHE WORD. BYTE ROTATE ACCESS AND MERGE WITH REGISTER 26. STORE INTO GENERAL REGISTER (S). SAVE ROTATED ACCESS IN REGISTER 26.
140
LEFT JUSTIFY STARTING BYTE. STORE INTO GENERAL REGISTER(S)
142
ANY BYTES=COMPARISON BYTE?
NO
144
COMPUTE NUMBER FOR MATCHING BYTE. UPDATE BYTE COUNT AND SET FLAG.
DECODE NEXT INSTRUCTION
159
146
END OF STRING IN THIS ACCESS?
NO
YES
148
ANY BYTES=COMPARISON BYTE?
YES
NO
150
INCREMENT GENERAL REGISTER POINTER. INCREMENT ADDRESS AND ACCESS NEXT CACHE WORD. BYTE ROTATE ACCESS AND MERGE WITH REGISTER 26. STORE INTO GENERAL REGISTER (S). SAVE ROTATED ACCESS IN REGISTER 26.
152
INCREMENT GENERAL REGISTER POINTER. STORE INTO GENERAL REGISTER(S). BYTE ROTATE ACCESS AND MERGE WITH REGISTER 26
154
ANY BYTES=COMPARISON BYTE?
NO
156
COMPUTE NUMBER FOR MATCHING BYTE. UPDATE BYTE COUNT AND SET FLAG.
158
DECODE NEXT INSTRUCTION

FIG. 8

ls rt, ra, rb
ss rt, ra, rb

FIG. 9A

```
Loop:   ls      Rw, Rs, Rx      /* Load 32 bytes from source */
        ss      Rw, Rd, Rx      /* Store 32 bytes to destination */
        add     Rx, Rx + 32     /* Increment Index */
        sub     Rcnt, Rcnt - 1  /* Decrement Count */
        brnz    Loop            /* Branch if Rcnt not zero */
```

FIG. 9B

```
Loop:   ls      Rv, Rs1, Rx     /* Load 32 bytes from source 1 */
        ls      Rw, Rs2, Rx     /* Load 32 bytes from source 2 */
        add     Rx, Rx + 16     /* Increment Index */
        cmp     Rv, Rw          /* Compare 1st Words */
        brne    Miscompare      /* Branch if not equal */
        cmp     Rv + 1, Rw + 1  /* Compare 2nd Words */
        brne    Miscompare      /* Branch if not equal */
        cmp     Rv + 2, Rw + 2  /* Compare 3rd Words */
        brne    Miscompare      /* Branch if not equal */
        cmp     Rv + 3, Rw + 3  /* Compare 4th Words */
        brne    Miscompare      /* Branch if not equal */
        sub     Rcnt, Rcnt - 1  /* Decrement Count */
        brnz    Loop            /* Branch if Rcnt not zero */
```

FIG. 9C

```
Loop:   lscb    Rw, Rs, Rx      /* Load up to 32 bytes from source */
        ss      Rw, Rd, Rx      /* Store same bytes to destination */
        add     Rx, Rx + 32     /* Increment Index */
        brne    Loop            /* Branch if null not found */
```

FIG. 9D